CHIA
center for health information and analysis

Application for Massachusetts All-Payer Claims Data (Non-Government)
[Exhibit A – Data Application]

I. INSTRUCTIONS

This form is required for all Applicants, Agencies, or Organizations, hereinafter referred to as "Organization", except Government Agencies as defined in 957 CMR 5.02, requesting protected health information. All Organizations must also complete the Data Management Plan, and attach it to this Application. The Application and the Data Management Plan must be signed by an authorized signatory. This Application and the Data Management Plan will be used by CHIA to determine whether the request meets the criteria for data release, pursuant to 957 CMR 5.00. Please complete the Application documents fully and accurately. Prior to receiving CHIA Data, the Organization must execute CHIA's Data Use Agreement. Organizations may wish to review that document prior to submitting this Application.

Before completing this Application, please review the data request information on CHIA's website:

- Data Availability
- Fee Schedule
- Data Request Process

After reviewing the information on the website and this Application, please contact CHIA at apcd.data@chiamass.gov if you have additional questions about how to complete this form.

The Appliciaton and all attachments must be uploaded to IRBNet. All Application documents can be found on the CHIA website.

Information submitted as part of the Application may be subject to verification during the review process or during any audit review conducted at CHIA's discretion.

Applications will not be reviewed until the Application and all supporting documents are complete and the required application fee is received.

A Fee Remittance Form with instructions for submitting the application fee is available on the CHIA website. If you are requesting a fee waiver, a copy of the Fee Remittance Form and any supporting documentation must be uploaded to IRBNet. Please be aware that if your research is funded and under that funding you are required to release raw data to the funding source, you may not receive CHIA Data.

II. FEE INFORMATION

1. Consult the most current Fee Schedule for All-Payer Claims Database data.

2. After reviewing the Fee Schedule, if you have any questions about the application or data fees, contact apcd.data@chiamass.gov.

3. If you believe that you qualify for a fee waiver, complete and submit the Fee Remittance Form and attach it and all required supporting documentation with your application. Refer to the Fee Schedule (effective Feb 1, 2017) for fee waiver criteria.

4. Applications will not be reviewed until the application fee is received.

5. Data for approved Applications will not be released until the payment for the Data is received.

III. ORGANIZATION & INVESTIGATOR INFORMATION

Project Title:	Assessment of Reach Out and Read's Early Literacy Program on Healthcare Cost and Utilization
IRBNet Number:	2320284-1
Organization Requesting Data (Recipient):	**Reach Out and Read, Inc**
Organization Website:	https://reachoutandread.org/
Authorized Signatory for Organization:	**Nikki Shearman**
Title:	Chief of Research & Innovation
E-Mail Address:	nikki.shearman@reachoutandread.org
Telephone Number:	857-636-1343
Address, City/Town, State, Zip Code:	89 South St, Suite 201, Boston, MA 02111
Data Custodian: **(individual responsible for organizing, storing, and archiving Data)**	**Rong Yi**
Title:	Principal, Data Science Modeling Practice
E-Mail Address:	rong.yi@milliman.com
Telephone Number:	781-462-8099
Address, City/Town, State, Zip Code:	1301 Fifth Avenue, Suite 3800 Seattle, WA 98101 USA
Primary Investigator (Applicant): **(individual responsible for the research team using the Data)**	**Rong Yi**
Title:	Principal, Data Science Modeling Practice
E-Mail Address:	rong.yi@milliman.com
Telephone Number:	781-462-8099
Address, City/Town, State, Zip Code:	1301 Fifth Avenue, Suite 3800 Seattle, WA 98101 USA
Names of Co-Investigators:	n/a
E-Mail Addresses of Co-Investigators:	n/a

IV. PROJECT INFORMATION

IMPORTANT NOTE: Organization represents that the statements made below as well as in any study or research protocol or project plan, or other documents submitted to CHIA in support of the Data Application are complete and accurate and represent the total use of the CHIA Data requested. Any and all CHIA Data released to the Organization under an approved application may ONLY be used for the express purposes identified in this section by the Organization, and for no other purposes. Use of CHIA Data for other purposes requires a separate Data Application to CHIA **or** written request to CHIA, with approval being subject to CHIA's regulatory restrictions and approval process. Unauthorized use is a material violation of your Organizations's Data Use Agreement with CHIA.

1. What will be the use of the CHIA Data requested? [Check all that apply]

☐ Epidemiological	☐ Health planning/resource allocation	☐Cost trends
☒ Longitudinal Research	☒ Quality of care assessment	☐ Rate setting
☐ Reference tool	☒ Research studies	☐ Severity index tool (or other derived input)
☐ Surveillance	☐ Student research	☒ Utilization review of resources

☐ Inclusion in a product ☐ Other (describe in box below)

Click here to enter text.

2. Provide an abstract or brief summary of the specific purpose and objectives of your Project. This description should include the research questions and/or hypotheses the project will attempt to address, or describe the intended product or report that will be derived from the requested data and how this product will be used. Include a brief summary of the pertinent literature with citations, if applicable.

This study aims to quantify the impact of Reach Out and Read's (ROR) early childhood literacy program on healthcare utilization and health outcomes in young children. ROR operates at the clinic level, integrating book distribution into pediatric care to promote early literacy and parent-child engagement. While existing literature supports the association between early literacy interventions and improved developmental and educational outcomes, this study seeks to evaluate whether ROR's program is associated with measurable differences in healthcare utilization patterns, including preventive care adherence, avoidable emergency department (ED) visits, hospital admissions, prescription adherence for chronic pediatric conditions, and early detection of developmental delays.

By integrating data from the Massachusetts All Payer Claims Database (APCD), geographically based socioeconomic data, and clinic-level data collected by ROR, we will analyze various healthcare cost, utilization, and quality outcomes among children in high-ROR-penetration regions and compare them to those in low-ROR-penetration regions. The study will employ both traditional statistical modeling and machine learning techniques to estimate differences in healthcare utilization while controlling for geographic, demographic, and socioeconomic factors. Our core hypothesis posits that children exposed to ROR's program demonstrate improved engagement with healthcare services, leading to increased preventive care uptake and better chronic disease management. The findings may provide evidence on whether ROR contributes to improved health outcomes, informing potential integration into value-based care models and healthcare policy discussions.

This research builds on prior studies indicating that early literacy interventions foster cognitive and socio-emotional development. To our knowledge, no previous studies have assessed the relationship between early literacy interventions and healthcare costs, utilization, and quality outcomes in the same manner. By providing empirical evidence linking ROR participation to quantifiable healthcare utilization metrics, we aim to contribute to the broader understanding of the impact of early literacy intervention programs.

The final product of this project will be a Milliman-issued white paper, commissioned by Reach Out and Read.

3. Has an Institutional Review Board (IRB) reviewed your Project?

☐ Yes [*If yes, a copy of the approval letter and protocol must be included with the Application package on IRBNet.*]
☒ No, this Project is not human subject research and does not require IRB review.

4. **Research Methodology**: Applications must include either the IRB protocol or a written description of the Project methodology (typically 1-2 pages), which should state the Project objectives and/or identify relevant research questions. This document must be included with the Application package on IRBNet and must provide sufficient detail to allow CHIA to understand how the Data will be used to meet objectives or address research questions.

V. PUBLIC INTEREST

1. Briefly explain why completing this Project is in the public interest. Use quantitative indicators of public health importance where possible, for example, numbers of deaths or incident cases; age-adjusted, age-specific, or crude rates; or years of potential life lost. *Uses that serve the public interest under CHIA regulations include, but are not limited to: health cost and utilization analysis to formulate public policy; studies that promote improvement in population health, health care quality or access; and health planning tied to evaluation or improvement of Massachusetts state government initiatives.*

This project aims to assess the impact of Reach Out and Read's (ROR) early childhood literacy program on healthcare utilization and health outcomes in Massachusetts, directly aligning with the public interest by informing health policy, improving population health, and potentially reducing healthcare costs.

Early childhood development is a critical determinant of long-term health outcomes. Research indicates that children who experience early literacy interventions demonstrate improved cognitive and socio-emotional development, which in turn may lead to better engagement with preventive healthcare services, improved chronic disease management, and reduced healthcare costs. Nationally, avoidable emergency department (ED) visits and hospitalizations for pediatric conditions such as asthma and preventable infections contribute significantly to healthcare expenditures. In Massachusetts, for example, avoidable ED visits among children account for approximately 20–25% of all pediatric ED visits annually, with Medicaid-insured children disproportionately represented in these preventable cases (see https://archives.lib.state.ma.us/server/api/core/bitstreams/15a837ac-b3c3-474d-980d-a335d64e5943/content). Studies show that early intervention and primary care engagement can significantly reduce these costly, avoidable episodes.

This project will use the Massachusetts All Payer Claims Database (APCD) to analyze the association between ROR exposure and key healthcare utilization metrics, such as:

- Preventive care adherence – increasing well-child visits and vaccinations, which are crucial for early disease prevention and developmental monitoring.
- Avoidable ED visits and hospital admissions – reducing unnecessary healthcare expenditures and improving care coordination.
- Prescription adherence for pediatric chronic conditions – particularly for asthma and Type 1 diabetes, which are leading causes of pediatric hospitalizations.
- Early diagnosis of developmental delays and disabilities – improving early intervention access and long-term educational and health outcomes.

By leveraging robust statistical and machine learning methodologies, the study may generate evidence on whether early literacy interventions at the clinic level contribute to measurable improvements in pediatric health outcomes. Findings from this study may guide statewide healthcare planning, value-based care initiatives, and Medicaid program improvements, particularly in promoting early childhood interventions that reduce long-term health costs and improve health equity.

In alignment with CHIA's mission, this research supports:

- Health cost and utilization analysis to inform public policy and healthcare payment reforms.
- Population health improvement by evaluating the impact of early literacy on healthcare access and quality.
- State health planning and evaluation of programs that intersect education, social determinants of health, and healthcare access for Massachusetts residents.

Given the potential for improved pediatric health outcomes and cost savings, this study serves the public interest by providing evidence for policies that enhance primary care, promote early literacy interventions, and reduce disparities in healthcare access and outcomes for young children across Massachusetts.

VI. DATASETS REQUESTED

The Massachusetts All-Payer Claims Database is comprised of medical, pharmacy, and dental claims and information from the member eligibility, provider, and product files that are collected from health insurance payers licensed to operate in the Commonwealth of Massachusetts. This information encompasses public and private payers as well as data from fully-insured and self-insured plans. APCD data are refreshed and updated annually and made available to approved data users. For more information about APCD Data, including available years of data and a full list of elements in the database please refer to layouts, data dictionaries and similar documentation included on CHIA's website.

Data requests are typically fulfilled on a one time basis, however; certain Projects may require future years of data that will become available in a subsequent release. Projects that anticipate a need for future years of data may request to be considered for a subscription. Approved subscriptions will receive, upon request, the same data files and data elements included in the initial Release annually or as available. Please note that approved subscription requests are subject to the Data Use Agreement, will require payment of fees for additional Data for Non-Government Entities, and subject to the limitation that the Data can be used only in support of the approved Project.

1. Please indicate below whether this is a one-time request, or if the described Project will require a subscription.

   ☒ One-Time Request **OR** ☐ Subscription

2. CHIA is currently supporting requests for claims data from 2016 to 2022. Requests made outside of these years may not be supported by CHIA and will be considered on a case-by-case basis. Please specify the years of data that are being requested: 2019 – 2023.

3. Specify below the data files requested for this Project, and provide your justification for requesting *each* file.

☒ **Medical Claims**
Describe how your research objectives require Medical Claims data: Medical claims are required to support: (1) identification of the five cost, utilization and quality outcome measures (2) identification of clinics that had implemented ROR's program (3) establishing a measure of relative health status at the individual level
☒ **Pharmacy Claims**
Describe how your research objectives require Pharmacy Claims data: Pharmacy claims are required to support the identification of Prescription refills for Type 1 Diabetes and Asthma
☐ **Dental Claims**
Describe how your research objectives require Dental Claims data: Click here to enter text.
☒ **Member Eligibility**
Describe how your research objectives require Member Eligibility data: Member eligibility data is required to support (1) identification of health insurance coverage types (2) merging of the county-level social determinants of health data from external sources (3) calculation of member months covered by a health insurer (4) inclusion and exclusion of members for modeling

☒ **Provider**
Describe how your research objectives require Provider data:
Provider data is required to support the identification of clinics and providers (provider NPI) that had implemented ROR's program
☐ **Product**
Describe how your research objectives require Product data:
Click here to enter text.

VII. DATA ENHANCEMENTS REQUESTED

State and federal privacy laws limit the release and use of CHIA Data to the minimum amount of data needed to accomplish a specific Project objective.

All-Payer Claims Database data is released in Limited Data Sets (LDS). All Organizations receive the "Core" LDS, but may also request the data enhancements listed below for inclusion in their analyses. Requests for enhancements will be reviewed by CHIA to determine whether each represents the minimum data necessary to complete the specific Project objective.

For a full list of elements in the release (i.e., the core elements and additional elements), please refer to release layouts, data dictionaries and similar documentation included on CHIA's website.

1. Specify below which enhancements you are requesting in addition to the "Core" LDS, provide your justification for requesting each enhancement.

a. Geographic Subdivisions

ZIP code and state geographic subdivisions are available for Massachusetts residents and providers only. Small population ZIP codes are combined with larger population ZIP codes. One ZIP Code per person (MEID) per year has been assigned based on the ZIP code/state reported in the member eligibility record's earliest submission year month. If the record does not have an MEID, assignment is based on distinct OrgID/Carrier Specific Unique Member ID.

Non-Massachusetts ZIP codes and state codes except for CT, MA, ME, NH, NY, RI, and VT are suppressed.

Select *one* of the following options.

☐ 3-Digit Zip Codes (standard)	☒ 5-Digit Zip Codes***
***If requested, provide justification for requesting 5-Digit Zip Code. Refer to specifics in your methodology: For inferring counties (imperfectly) at the member level, so that we may be able to bring in county-level social determinants of health data.	

b. Date Resolution

Select *one* option from the following options.

☐ Year (YYYY) (Standard)	☐ Month (YYYYMM) ***	☒ Day (YYYYMMDD) *** [for selected data elements only]

***** If requested, provide justification for requesting Month or Day. Refer to specifics in your methodology:**

As shown below, the outcomes we are modeling require precision at the YYYYMMDD level. These healthcare utilization counts all require specific and precise service dates, admission and discharge dates at the claim line level in calculations.

- Preventive care visits
- Avoidable Emergency Department visits
- Avoidable hospital admissions and readmissions
- Prescription refills for Type 1 Diabetes and Asthma
- Age at initial diagnosis of developmental delays or developmental disabilities

c. National Provider Identifier (NPI)

Select *one* of the following options.

☐ Encrypted National Provider Identifiers (standard)	☒ Decrypted National Provider Identifiers***

***** If requested, provide justification for requesting decrypted National Provider Identifier(s). Refer to specifics in your methodology:**

Decrypted NPI allows us to join the provider data and medical claims data to an external database of providers and health systems, which can be used to identify providers (NPIs) that were affiliated with clinics that had implemented ROR's program.

VIII. MEDICAID (MASSHEALTH) DATA

1. Please indicate whether you are seeking Medicaid Data:

☒ Yes
☐ No

2. Federal law (42 USC 1396a(a)7) restricts the use of individually identifiable data of Medicaid recipients to uses that are ***directly connected to the administration of the Medicaid program***. If you are requesting MassHealth Data, please describe, in the space below, why your use of the Data meets this requirement. *Your description should focus on how the results of your project could be used by the Executive Office of Health and Human Services in connection with the administering the MassHealth program*. Requests for identifiable MassHealth Data will be forwarded to MassHealth for a determination as to whether the proposed use of the Data is directly connected to the administration of the MassHealth program. CHIA cannot release MassHealth Data without approval from MassHealth. This may introduce significant delays in the receipt of MassHealth Data.

This project aims to quantify the associations between ROR's early childhood literacy program implementation with 5 healthcare utilization and outcome measures. We also include a short discussion on the relevance of these measures to MassHealth.

- Preventive care visits

  Regular well-child visits, vaccinations, and screenings for developmental, vision, hearing, and chronic conditions enable early detection and management, preventing complications and costly acute care such as ER visits or hospitalizations. While

preventive services may raise short-term costs, they often yield long-term savings for Medicaid by avoiding complex treatments.

- Avoidable Emergency Department visits

  For children ages 0–5 on Medicaid, many emergency department (ED) visits arise from conditions that could be prevented or managed in primary care or urgent care settings, such as fevers, minor injuries, or mild asthma exacerbations. Timely access to well-child visits, same-day sick care, and effective chronic condition management—combined with caregiver education on when and where to seek care—can significantly reduce unnecessary ED use.

- Avoidable hospital admissions and readmissions

  For children ages 0–5 on Medicaid, avoidable hospital admissions and readmissions often stem from preventable complications of respiratory infections, asthma, dehydration, or other common pediatric conditions. Strengthening primary care access, ensuring timely follow-up after ED or urgent care visits, and supporting families with clear care plans can prevent many inpatient stays.

- Prescription refills for Type 1 Diabetes and Asthma

  Type 1 Diabetes and asthma are prevalent conditions amongst children receiving Medicaid benefits. Medication adherence is vital for improving health outcomes and reducing avoidable healthcare services. Noncompliance or inconsistent compliance often leads to exacerbations which could result in ED visits, hospital admissions or worse.

- Age at initial diagnosis of developmental delays or developmental disabilities

  In Massachusetts, 32% children with special needs are covered by MassHealth (https://www.kff.org/medicaid/issue-brief/5-key-facts-about-children-with-special-health-care-needs-and-medicaid/#Appendix). Children with special needs typically have higher healthcare expenditures than those without. There is substantial empirical evidence that earlier diagnosis and intervention would lead to better health outcomes and lower healthcare utilizations down the road for children with special needs.

For all of the above five measures, our hypothesis is that ROR's integration into pediatric visits provides a trusted touchpoint to strengthen caregiver engagement, reinforce health education, and promote consistent primary care use, better chronic disease management, earlier identification of developmental issues. If there is empirial evidence to support the hypothesis, MassHealth may consider encouraging adoption of ROR's program for young beneficiaries.

This research directly aligns with MassHealth's mission to improve healthcare quality, reduce unnecessary costs, and ensure equitable access to care for Medicaid-enrolled children. By leveraging MassHealth claims data, this study will provide valuable insights into how early literacy interventions contribute to better health outcomes, informing future Medicaid strategies and investments in pediatric care.

1. How does the project relate directly to the administration of the Medicaid program?

In Massachusetts, approximately 40% of children are enrolled in MassHealth. Given that this study focuses on children aged 0-5—an age group with particularly high Medicaid enrollment—MassHealth claims data is essential for a comprehensive analysis. Additionally, based on ROR's experience, its programs are more prevalent in lower socioeconomic areas where Medicaid participation is higher. Without access to Medicaid claims, we would be unable to fully capture the impact of ROR's program on its target population.

By analyzing MassHealth claims data, this study will assess whether exposure to ROR is associated with improved preventive care adherence, reduced avoidable emergency department (ED) visits and hospitalizations, improved chronic disease management, and earlier detection of developmental delays. Understanding these associations may help inform Medicaid policy regarding early childhood interventions that improve long-term health outcomes and reduce unnecessary healthcare spending.

2. What specific Medicaid program, policy, rule, or law will be affected or changed based on the outcome of this project?

Findings from this study may have the potential to impact several Medicaid policies and programs, including:

- If the study finds that ROR participation is associated with improved preventive care and reduced avoidable ED visits, Medicaid could consider incentivizing its integration into pediatric practices participating in managed care and value-based payment structures (MassHealth ACOs)
- The results could support initiatives that encourage early literacy interventions as a tool for improving pediatric health outcomes, aligning with the broader goals of Medicaid's Early and Periodic Screening, Diagnostic, and Treatment (EPSDT) program.

3. How will MassHealth's objectives be helped or impaired by approving this project?

We do not foresee negative impacts on MassHealth. The study will use existing claims data and does not interfere with Medicaid operations or service delivery. On the other hand, this project may support MassHealth's objectives by:

- Providing data-driven insights into the impact of early childhood literacy on healthcare utilization and costs, helping to inform Medicaid program design and policy decisions.
- Identifying opportunities to enhance preventive care engagement within pediatric practices serving Medicaid-enrolled children.
- Assessing healthcare utilization patterns and disparities among Medicaid beneficiaries to support population health improvement efforts.

4. Will the results of the research have the potential for:

- **Reducing costs of the Medicaid program?** Yes. If ROR is associated with reduced avoidable ED visits and hospitalizations, as well as improved chronic disease management, it could lead to cost savings for MassHealth. Avoidable pediatric ED visits account for a substantial portion of Medicaid expenditures, and interventions that improve primary care engagement can help mitigate unnecessary spending.
- **Improving access for recipients?** Yes. If literacy interventions improve parent engagement with healthcare services, more children may receive timely well-child visits, immunizations, and screenings, improving access to necessary preventive care.
- **Increasing quality of care to recipients?** Yes. Early identification of developmental delays, better chronic disease management, and increased preventive care adherence all contribute to improved long-term health outcomes, aligning with Medicaid's quality improvement goals.

5. Project Deliverables for MassHealth

The final project report will be provided to ROR as a Milliman white paper commissioned by ROR ready for broad distribution. With ROR's approval, the report will be made available to MassHealth, as well as the general public.

6. How MassHealth Can Use the Project Deliverables

MassHealth can use the study findings to:

- Support evidence-based policy decisions regarding pediatric preventive care and early childhood interventions.
- Assess cost-effectiveness and return on investment of integrating literacy-based programs into Medicaid's value-based care strategies.
- Enhance reimbursement strategies by identifying whether ROR participation improves long-term health outcomes and reduces high-cost utilization, thereby informing Medicaid managed care contracts and incentive structures.
- Strengthen Medicaid's pediatric care framework by supporting programs that enhance early childhood development and reduce disparities in healthcare access.

3. Organizations approved to receive Medicaid Data will be required to execute a Medicaid Acknowlegment of Conditions. MassHealth may impose additional requirements on applicants for Medicaid Data as necessary to ensure compliance with federal laws and regulations regarding Medicaid.

IX. DATA LINKAGE

Data linkage involves combining CHIA Data with other data to create a more extensive database for analysis. Data linkage is typically used to link multiple events or characteristics within one database that refer to a single person within CHIA Data.

1. Do you intend to link or merge CHIA Data to other data?

☒ Yes

☐ No linkage or merger with any other data will occur

2. If yes, please indicate below the types of data to which CHIA Data will be linked. [Check all that apply]

☐ Individual Patient Level Data (e.g. disease registries, death data)

☒ Individual Provider Level Data (e.g., American Medical Association Physician Masterfile)

☐ Individual Facility Level Data (e.g., American Hospital Association data)

☒ Aggregate Data (e.g., Census data)

☒ Other (please describe): Milliman's Provider Registry Data https://medinsight.com/healthcare-data-analytics-software/data-intelligence/provider-registry/

3. If yes, describe the dataset(s) to which the CHIA Data will be linked, indicate which CHIA Data elements will be linked and the purpose for each linkage.

We plan to (1) Merge the CDC's Social Vulnerability Index ((https://www.atsdr.cdc.gov/place-health/php/svi/index.html) and its underlying census measures with the CHIA data at the county level, which would allow for the models to control for SDOH differences in healthcare outcomes. The SVI includes key socioeconomic indicators such as income level, education, housing conditions, and access to transportation, all of which influence healthcare access and utilization. By incorporating these factors, we aim to account for the potential impact of environmental and community-level disadvantages on health outcomes among children in the study. (2) Provide ROR's TIN-to-clinic list to CHIA with a starting year indicator so that clinics that had implemented ROR's early literacy programs may be identified, along with the year that the clinics started implementation. We would ask that CHIA includes two additional fields to indicate these clinics and their starting years. (3) Use Milliman's Provider Registry Data (https://medinsight.com/healthcare-data-analytics-software/data-intelligence/provider-registry/) to further identify individual providers working at the clinics that had implemented ROR's programs. Milliman's Provider Registry Data is a comprehensive, continually updated database that tracks provider affiliations, specialties, and locations, ensuring accurate provider attribution. By linking these datasets, we will enhance the identification of ROR-participating clinics and providers, improving the precision of our attribution model and strengthening the study's analytical rigor.

4. If yes, for each proposed linkage above, please describe your method or selected algorithm (e.g., deterministic or probabilistic) for linking each dataset. If you intend to develop a unique algorithm, please describe how it will link each dataset.

The data linkages proposed in the above are deterministic and straightforward and will not involve any probabilistic matching. Specifically,

- We will use the 5-digit zip code in the Member Eligibility data to create counties. We understand there are limitations to this method. After creating counties, we will merge the CDC's SVI (https://www.atsdr.cdc.gov/place-health/php/svi/index.html) data with the CHIA data by county.
- We will use the clinics TIN from ROR, the NPI-to-TIN mapping from the Milliman Provider Registry data to crosswalk to the TINs on CHIA's Provider Data. Depending on data quality, some manual mapping may be needed.

5. If yes, attach or provide below a complete listing of the variables from all sources to be included in the final linked analytic file.

The CDC SVI data is in the public domain and can be retrieved here: https://www.atsdr.cdc.gov/place-health/php/svi/index.html.
The Milliman Provider Registry data is proprietary to Milliman, but will include TINs and NPIs among other provider and health system information.

6. If yes, please identify the specific steps you will take to prevent the identification of individual patients in the linked dataset.

Milliman is the custodian of the CHIA data extract. Milliman is a HITRUST entity.

The data files we request from CHIA limited datasets.

To ensure compliance with privacy regulations and maintain the confidentiality of individual patients, we will implement the following measures when linking and analyzing the dataset:

Indirect identifiers, such as dates of birth, service dates, and rare conditions, will either be generalized (e.g., age in years instead of birthdate) or grouped into broad disease categories and rolled up to risk scores to prevent linkage to specific individuals. Patient location data will be aggregated at the county level. Any cell sizes less than 11 individuals will be combined with other cells/groups by demographic, geographic or health status categories.

Only authorized personnel will have access to the datasets in Milliman's secure environment. No direct extraction is feasible.

X. PUBLICATION / DISSEMINATION / RE-RELEASE

1. Do you anticipate that the results of your analysis will be published or made publically available? If so, how do you intend to disseminate the results of the study (e.g.; publication in professional journal, poster presentation, newsletter, web page, seminar, conference, statistical tabulation)? Any and all publication of CHIA Data must comply with CHIA's cell size suppression policy, as set forth in the Data Use Agreement. Please explain how you will ensure that any publications ***will not disclose a cell less than 11***, and percentages or other mathematical formulas that result in the display of a cell less than 11.

The main work product is a Milliman-issued white paper, commissioned by Reach Out and Read (ROR), summarizing key findings for policymakers and stakeholders. The findings may be presented by either Milliman with ROR's permission, or by ROR, at conferences and circulated in newsletters and websites. The results will be reviewed for sample size considerations prior to publication.

2. Describe your plans to use or otherwise disclose CHIA Data, or any Data derived or extracted from such Data, in any paper, report, website, statistical tabulation, seminar, or other setting that is not disseminated to the public.

Processed datasets, such as risk scores and HCG service categories, the five outcome measures are all derived from the CHIA data extracts. They will be stored in Milliman's HIPAA compliant secure environment for up to two years after completion of the project, and will be archived for an additional three years before deletion. Milliman will not use the CHIA data extracts or the processed datasets derived from the CHIA data extracts for purposes other than for the said ROR program impact assessment project.

3. What will be the lowest geographical level of analysis of data you expect to present for publication or presentation (e.g., state level, city/town level, zip code level, etc.)? Will maps be presented? If so, what methods will be used to ensure that individuals cannot be identified?

The lowest geographic level for this project is county. We will not show maps.

4. Will you be using CHIA Data for consulting purposes?
- ☐ Yes
- ☒ No

5. Will you be selling standard report products using CHIA Data?
- ☐ Yes
- ☒ No

6. Will you be selling a software product using CHIA Data?
- ☐ Yes
- ☒ No

7. Will you be using CHIA Data as in input to develop a product (i.e., severity index took, risk adjustment tool, reference tool, etc.)
- ☐ Yes
- ☒ No

8. Will you be reselling CHIA Data in any format not noted above?

☐ Yes
☒ No

If yes, in what format will you be reselling CHIA Data?

Click here to enter text.

9. If you have answered "yes" to questions 5, 6, 7 or 8, please provide the name and a description of the products, software, services, or tools.

Click here to enter text.

10. If you have answered "yes" to questions 5, 6, 7 or 8, what is the fee you will charge for such products, software, services or tools?

Click here to enter text.

XI. APPLICANT QUALIFICATIONS

1. Describe your previous experience using claims data. This question should be answered by the primary investigator and any co-investigators who will be using the Data.

Milliman is the primary investigator for the said ROR program impact assessment project. Milliman is a leading healthcare analytics and actuarial consulting firm. Most of Milliman Health Discipline's work involves in using claims data. The lead consultant, Rong Yi, not only has more than 20 years of experience using claims data, but also has been deeply involved in using APCD data, including the MA APCD. She led the development, implementation and operation of the Massachusetts ACA risk adjustment program between 2012 and 2017, during which she worked closely with CHIA on claims data submissions and risk adjustment funds settlement calculations. The claim-based models and methodologies she developed have been widely used in the industry both in the US and in other parts of the world.

2. **Resumes/CVs**: When submitting your Application package on IRBNet, include résumés or curricula vitae of the principal investigator and co-investigators. (These attachments will not be posted on the internet.)

XII. USE OF AGENTS AND/OR CONTRACTORS

By signing this Application, the Organization assumes all responsibility for the use, security and maintenance of the CHIA Data by its agents, including but not limited to contractors. The Organization must have a written agreement with the agent of contractor limiting the use of CHIA Data to the use approved under this Application as well as the privacy and security standards set forth in the Data Use

Agreement. CHIA Data may not be shared with any third party without prior written consent from CHIA, or an amendment to this Application. CHIA may audit any entity with access to CHIA Data.

Provide the following information for **all** agents and contractors who will have access to the CHIA Data. [*Add agents or contractors as needed.]*

AGENT/CONTRACTOR #1 INFORMATION	
Company Name:	**Milliman, Inc.**
Company Website	https://us.milliman.com/en
Contact Person:	**Rong Yi**
Title:	Principal, Data Science Modeling Practice
E-mail Address:	rong.yi@milliman.com
Address, City/Town, State, Zip Code:	1301 Fifth Avenue, Suite 3800 Seattle, WA 98101 USA
Telephone Number:	781-462-8099
Term of Contract:	February 10, 2025 -

1. Describe the tasks and products assigned to the agent or contractor for this Project and their qualifications for completing the tasks.

Milliman will oversee data acquisition, study design, analysis, and reporting to assess the impact of Reach Out and Read (ROR) on healthcare utilization. The firm will first assist in obtaining Massachusetts APCD data from CHIA and merge it with the CDC Social Vulnerability Index (SVI) and Milliman's Provider Registry Data to enhance geographic and provider-level analysis. Ensuring compliance with HIPAA and CHIA's data security policies, Milliman will handle secure data integration and management.

In the study design phase, Milliman will define the study population, constructing a matched comparison between children in high- and low-ROR-penetration regions. The team will apply statistical and machine learning models to estimate ROR's impact on healthcare utilization while controlling for geographic, demographic, and socioeconomic factors. Risk adjustment methods will be incorporated to account for differences in health status across the study population.

For outcome measurement, Milliman will analyze differences in preventive care adherence, avoidable emergency department visits, hospitalizations, prescription adherence, and early diagnosis of developmental delays. To ensure compliance with CHIA's cell suppression policy, the team will apply de-identification and small-N suppression techniques, preventing the reporting of any identifiable patient information.

Finally, Milliman will produce a technical report summarizing methodology and findings, as well as a Milliman-issued white paper commissioned by ROR. These results will be presented in policy briefings for MassHealth and disseminated through conference presentations, professional journals, and web-based materials.
Milliman's Qualifications
Milliman brings deep expertise in healthcare data analytics and Medicaid research, with a team of actuaries, health economists, clinicians, and data scientists:

- Dr. Rong Yi, PhD – Principal, oversees study design and modeling strategy.
- Dr. Tzu-Chun Kuo, PhD – Senior Health Economist, leads statistical analysis.
- Dr. David Mirkin, MD – Chief Medical Officer, ensures clinical validity.
- Meseret Woldeyes – Managing Data Scientist, develops machine learning models.
- Robert Richards, PhD – Senior Manager, leads provider data integration.

- Mia Wafer – GIS Analyst, integrates SVI data for geographic analysis.

2. Describe the Organization's oversight and monitoring of the activities and actions of the agent or contractor for this Project, including how the Organization will ensure the security of the CHIA Data to which the agent or contractor has access.

ROR will conduct frequent status meetings with Milliman to review project progress, data handling procedures, and compliance with CHIA security policies. Milliman will report to ROR and to CHIA any potential risks or issues related to data integrity and confidentiality.

Prior to any external dissemination of study findings, ROR will review all publications, reports, and presentations to ensure compliance with CHIA's cell size suppression policy and other confidentiality requirements. If necessary, study results will be submitted for CHIA review and approval before public release.

By maintaining continuous oversight, strict data security measures, and clear accountability mechanisms, ROR will ensure that CHIA data is used appropriately, securely, and in full compliance with regulatory requirements.

3. Will the agent or contractor have access to and store the CHIA Data at a location other than the Organization's location, off-site server and/or database?

☒ Yes
☐ No

4. If yes, a separate Data Management Plan **must** be completed by the agent or contractor.

AGENT/CONTRACTOR #1 INFORMATION	
Company Name:	**Milliman, Inc.**
Company Website	https://us.milliman.com/en
Contact Person:	**Rong Yi**
Title:	Principal, Data Science Modeling Practice
E-mail Address:	rong.yi@milliman.com
Address, City/Town, State, Zip Code:	1301 Fifth Avenue, Suite 3800 Seattle, WA 98101 USA
Telephone Number:	781-462-8099
Term of Contract:	February 10, 2025 -

1. Describe the tasks and products assigned to the agent or contractor for this Project and their qualifications for completing the tasks.

Milliman will oversee data acquisition, study design, analysis, and reporting to assess the impact of Reach Out and Read (ROR) on healthcare utilization. The firm will first assist in obtaining Massachusetts APCD data from CHIA and merge it with the CDC Social Vulnerability Index (SVI) and Milliman's Provider Registry Data to enhance geographic and provider-level analysis. Ensuring compliance with HIPAA and CHIA's data security policies, Milliman will handle secure data integration and management.

In the study design phase, Milliman will define the study population, constructing a matched comparison between children in high- and low-ROR-penetration regions. The team will apply statistical and machine learning models to estimate ROR's impact on healthcare utilization while controlling for geographic, demographic, and socioeconomic factors. Risk adjustment methods will be incorporated to account for differences in health status across the study population.

For outcome measurement, Milliman will analyze differences in preventive care adherence, avoidable emergency department visits, hospitalizations, prescription adherence, and early diagnosis of developmental delays. To ensure compliance with CHIA's cell suppression policy, the team will apply de-identification and small-N suppression techniques, preventing the reporting of any identifiable patient information.

Finally, Milliman will produce a technical report summarizing methodology and findings, as well as a Milliman-issued white paper commissioned by ROR. These results will be presented in policy briefings for MassHealth and disseminated through conference presentations, professional journals, and web-based materials.

Milliman's Qualifications

Milliman brings deep expertise in healthcare data analytics and Medicaid research, with a team of actuaries, health economists, clinicians, and data scientists:

- Dr. Rong Yi, PhD – Principal, oversees study design and modeling strategy.
- Dr. Tzu-Chun Kuo, PhD – Senior Health Economist, leads statistical analysis.
- Dr. David Mirkin, MD – Chief Medical Officer, ensures clinical validity.
- Meseret Woldeyes – Managing Data Scientist, develops machine learning models.
- Robert Richards, PhD – Senior Manager, leads provider data integration.
- Mia Wafer – GIS Analyst, integrates SVI data for geographic analysis.

2. Describe the Organization's oversight and monitoring of the activities and actions of the agent or contractor for this Project, including how the Organization will ensure the security of the CHIA Data to which the agent or contractor has access.

ROR will conduct frequent status meetings with Milliman to review project progress, data handling procedures, and compliance with CHIA security policies. Milliman will report to ROR and to CHIA any potential risks or issues related to data integrity and confidentiality.

Prior to any external dissemination of study findings, ROR will review all publications, reports, and presentations to ensure compliance with CHIA's cell size suppression policy and other confidentiality requirements. If necessary, study results will be submitted for CHIA review and approval before public release.

By maintaining continuous oversight, strict data security measures, and clear accountability mechanisms, ROR will ensure that CHIA data is used appropriately, securely, and in full compliance with regulatory requirements.

3. Will the agent or contractor have access to or store the CHIA Data at a location other than the Organization's location, off-site server and/or database?

☐ Yes
☒ No

4. If yes, a separate Data Management Plan **must** be completed by the agent or contractor.

[INSERT A NEW SECTION FOR ADDITIONAL AGENTS/CONTRACTORS AS NEEDED]

XIII. ATTESTATION

By submitting this Application, the Organization attests that it is aware of its data use, privacy and security obligations imposed by state and federal law *and* confirms that it is compliant with such use, privacy and security standards. The Organization further agrees and understands that it is solely responsible for any breaches or unauthorized access, disclosure or use of CHIA Data, including, but not limited to, any breach or unauthorized access, disclosure or use by any third party to which it grants access.

Organizations approved to receive CHIA Data will be provided with Data following the payment of applicable fees and upon the execution of a Data Use Agreement requiring the Organization to adhere to processes and procedures designed to prevent unauthorized access, disclosure or use of data.

By my signature below, I attest: (1) to the accuracy of the information provided herein; (2) this research is not funded by a source requiring the release of raw data to that source; (3) that the requested Data is the minimum necessary to accomplish the purposes described herein; (4) that the Organization will meet the data privacy and security requirements described in this Application and supporting documents, and will ensure that any third party with access to the Data meets the data use, privacy and security requirements; and (5) to my authority to bind the Organization.

Signature: (Authorized Signatory for Organization)	
Printed Name:	**Rong Yi**
Title:	Principal, Data Science Modeling Practice
Date:	08/20/2025

Attachments:
A completed Application must have the following documents attached to the Application or uploaded separately to IRBNet:
☐ 1. IRB approval letter and protocol (if applicable), or research methodology (if protocol is not attached)
☒ 2. Data Management Plan (including one for each agent or contractor that will have access to or store the CHIA Data at a location other than the Organization's location, off-site server and/or database);
☒ 3. CVs of Investigators (upload to IRBNet)

APPLICATIONS WILL NOT BE REVIEWED UNTIL THEY ARE COMPLETE, INCLUDING ALL ATTACHMENTS.